J. Statist. Comput. Simul., Vol. 34, pp. 145–154
Reprints available directly from the publisher
Photocopying permitted by license only



# AN APPROACH TO ESTIMATING DENSITY FROM LINE TRANSECT DATA WHERE THE ANIMALS MOVE IN RESPONSE TO THE OBSERVER

RICHARD M. ENGEMAN and JEFFREY F. BROMAGHIN

*Denver Wildlife Research Center, Animal Damage Control Program, APHIS, USDA, Bldg. 16, Denver Federal Center, Denver, Colorado 80225-0266, USA*

*(Received 2 June 1988; in final form 28 August 1989)*

An approach is presented for estimating density from line transect data where the animals move away in response to the observer. This estimation method involves fitting the perpendicular sighting distances to the Weibull distribution. Estimates of the Weibull parameters are incorporated into the density estimate. When movement exists, even light movement, this method greatly outperforms the Fourier series estimator. It also outperforms the Fourier series estimator in some low density situations when there is no movement. The Weibull method shows promise for estimating density in the presence of movement and further development is planned.

## 1. INTRODUCTION

Line transect sampling is commonly used as a practical means for estimating the density of objects in an area. Current analytical methods for data from line transect sampling assume that the objects of interest do not move in response to the observer. However, if the objects are animals, it is not uncommon for them to move in response to the observer prior to being observed. Usually the response is to move away from the observer, but less commonly they can be attracted towards the observer. We concern ourselves only with the more frequently encountered situation where the animals are moving away from the observer (and hence the transect line). In the presence of movement, perpendicular distance measurements from the line to the animal tend to be longer than if the animal had not responded to the observer.

We briefly describe density estimation using perpendicular distance data from the line transect sampling. Burnham, Anderson and Laake (1980) present a comprehensive discussion of density estimation from line transect sampling. Because that publication is one of the standard references for line transect sampling, we follow their notation closely. We present a possible approach to density estimation when there is movement in response to the observer and examine its potential through a simulation study that compares it to the Fourier Series estimator (Crain *et al.*, 1978).

## 2. ESTIMATION WITHOUT MOVEMENT

When measuring perpendicular distances to objects from a line transect, the expected number of objects detected in a strip of width $w$ (on either side of the transect), and length $L$ is

$$E(n) = D2wLP_w, \tag{1}$$

where $D$ is the density of objects and $P_w$ is the unconditional average detection probability. An estimate of density is (Burnham *et al.*, 1980)

$$\hat{D} = \frac{n}{2wLP_w}. \tag{2}$$

In the absence of movement in response to the observer, $P_w$ in Eq. (2) is defined as

$$P_w = 1/w \int_0^w g(x)\,dx, \tag{3}$$

where $g(x)$ is the sighting function describing the detectability of the objects at distance $x$ from the transect line. The $1/w$ term arises from the assumption that, without movement, if all objects were sighted, then their distances from the randomly placed transect would be distributed uniformly on the interval $(0, w)$. The sighting function, $g(x)$, is considered to be monotonically nonincreasing with the assumption that $g(0) = 1$. This assumption implies that all objects that fall directly on the line are detected with probability 1. A probability density function ($pdf$), $f(x)$, can be defined as

$$f(x) = \frac{1/wg(x)}{1/w \int_0^w g(x)\,dx} = \frac{g(x)}{\int_0^w g(x)\,dx}. \tag{4}$$

The assumption that $g(0) = 1$ implies that

$$f(0) = \frac{1}{\int_0^w g(x)\,dx},$$

or

$$\int_0^w g(x)\,dx = \frac{1}{f(0)}. \tag{5}$$

After estimating $f(0)$ based on the observed perpendicular distances and substituting Eqs. (5) and (3) into Eq. (2), we have the estimate of density

$$\hat{D} = \frac{n\hat{f}(0)}{2L}. \tag{6}$$

Many methods have been proposed for estimating $f(0)$ (e.g., see Burnham *et al.*, 1980 and Seber, 1986). The Fourier series estimator of Crain *et al.* (1978) is one of the most flexible and widely accepted methods (Burnham *et al.*, 1980 and Seber, 1986).

## 3. AN APPROACH TO THE MOVEMENT PROBLEM

If animals move away from the line in response to the observer, then the conditional distribution of their distances to the randomly placed transect is no longer uniform and Eq. (3) is rewritten as (see Burnham *et al.*, 1980)

$$P_{mw} = \int_0^w g(x)h(x \mid w)\,dx, \tag{7}$$

where $h(x \mid w)$ is the conditional *pdf* of perpendicular distances after movement and $m$ used as a subscript denotes movement. Now $f(x)$ is related to $P_{mw}$ by

$$f_m(x) = \frac{g(x)h(x \mid w)}{\int_0^w g(x)h(x \mid w)\,dx} = kg(x)h(x \mid w), \tag{8}$$

where $k$ is a constant denoting the reciprocal of the definite integral in the denominator. An estimate of $P_{mw}$ can no longer be found by estimating $f(0)$.

It may not be unreasonable to assume that if animals are moving away from the line in response to the observer, then the probability of observing an animal on the line is 0. The assumption that $g(0)=1$ is maintained, but $h(0 \mid w)=0$ and

$$f_m(0) = \frac{g(0)h(0 \mid w)}{\int_0^w g(x)h(x \mid w)\,dx} = 0. \tag{9}$$

Obviously, the *pdf* in Eq. (8) should be of a functional form that fits the observed perpendicular distance data where movement has occurred. One candidate is the *pdf* for the Weibull distribution:

$$f_w(x) = \frac{b}{a^b} x^{b-1} \exp\left(-\left(\frac{x}{a}\right)^b\right), \tag{10}$$

where $x>0$, $a>0$, $b>0$. It should be noted that Ramsey (1979) discussed a "detectability curve kernel" *based* on the Weibull distribution (also described by Pollock, 1978, as the exponential power family), however the observed distances did not follow the Weibull distribution, nor was the issue of animal movement considered.

We next define $f_w(x)$ in terms of Eq. (8). If we let

$$g(x)=\exp\left(-\left(\frac{x}{a}\right)^{b}\right), \tag{11}$$

then $g(0)=1$, and $g(x)$ is an exponential decay, which was used by Gates *et al.* (1968) for a sighting function in the first probabilistic development of line transect analyses. We then let

$$h(x \mid w)=x^{b-1} \tag{12}$$

which allows $h(0 \mid w)=0$.

Consider now that Eq. (10) is of the form given in Eq. (8), that is

$$\frac{b}{a^{b}}x^{b-1}\exp\left(-\left(\frac{x}{a}\right)^{b}\right)=g(x)h(x \mid w)\frac{1}{\int_0^w g(x)h(x \mid w)\,dx}. \tag{13}$$

We defined $g(x)=\exp(-(x/a)^{b})$ and $h(x \mid w)=x^{b-1}$, leaving

$$\frac{b}{a^{b}}=\frac{1}{\int_0^w g(x)h(x \mid w)\,dx}=\frac{1}{p_{mw}}. \tag{14}$$

Substituting into Eq. (2) we have

$$\hat{D}=\frac{n}{2wL}\cdot\frac{1}{(b/a^{b})}=\frac{na^{b}}{2wLb} \tag{15}$$

where $a$ and $b$ are the Weibull parameters to be estimated using observed perpendicular distances.

## 4. SIMULATION STUDY

The performance of the method described in the preceding section (which we refer to as the Weibull method) was examined through a Monte Carlo simulation study. The Fourier series method for estimating density in Eq. (6) was used as a standard of comparison due to its widespread acceptance (Seber, 1986) as one of the "preferred" estimators (e.g., it is the default method for program TRANSECT; Laake, Burnham and Anderson, 1979).

### *4.2 Simulation Setup*

We designed a simulation study to examine the performance of the Weibull and Fourier estimators over a range of population densities with several levels of movement in response to an observer. We considered two shapes for the sighting function with three parameterizations of each to represent different sighting abilities.

Three densities of objects were considered for computer simulation: 2, 4, 8. For a specified area of width $W$ and length $L$, the number of objects needed to achieve a density $D$ is $n = 2\ WLD$. Using the BASIC internal random number generator (TRS-XENIX operating system MBASIC), a population of objects (coordinates) was generated for the specified density, length and width of a survey area. A new seed was supplied for each random number generated. Previous extensive testing by the authors indicated close conformation of the generated numbers to a uniform (0, 1) distribution. Multiplying the uniform (0, 1) numbers by $W$ produces object distances located such that their distances from the transect line were distributed uniformly over the interval $(0, W)$. Note that because the estimation methods are not concerned with which side of the transect line an object occurs, we need only have the $n$ objects distributed over $(0, W)$ but the total width is still $2W$ for estimation purposes. A new population of objects was created at each iteration of a simulation run.

At each iteration, a sample of objects was randomly selected by applying one of two sighting functions. The first sighting function resembles many of the data sets we have seen in practice (where no observer related movement is detectable) and is based on the following function

$$\exp(-x^2/p_1), \tag{16}$$

where $x$ is the perpendicular distance from the transect and $p_1$ is a user specified parameter that determines sighting ability. Thus, the probability that an object will be included in the sample follows the shape of half of a normal distribution. Three values for $p_1$ were selected for simulating a variety of sighting abilities: 0.5, 1.5, and 3.0.

An alternate shape for the sighting function which also resembles field data is the following exponential decay

$$\exp(-x/p_2), \tag{17}$$

where $x$ again is the perpendicular distance to the transect and $p_2$ is a user specified parameter that determines sighting ability. Three values were also considered for $p_2$; 0.5, 0.75, and 1.0.

Movement was incorporated by adding to the objects' original distance from the transect line according to a movement function. We decided that a reasonable movement function would be one where the objects closest to the observer (transect line) would move the furthest. We also felt it reasonable to base the movement on the functional form of the normal distribution. Thus, to model movement we added to the object's distance from the transect according to the function

$$\frac{\exp(-x^2/p_3)}{p_3\sqrt{2\pi}}, \tag{18}$$

where $x$ is the original perpendicular distance to the transect line and $p_3$ is a parameter that determines the extremity of movement. For $p_3$, we considered the values 0.25, 0.50, 1.0 and 2.0. The values 0.25 and 0.50 represent extreme movement, 1.0 represents moderate movement, and 2.0 represents light movement.

The Fourier series estimates were calculated according to the formulae and methods in Burnham *et al.* (1980). There are many available methods for estimating Weibull parameters (e.g., see Engeman and Keefe, 1982, 1985). Maximum likelihood was used to estimate the Weibull density parameters because it performs well (Engeman and Keefe, 1982, 1985) and is widely accepted. The formulae and iterative algorithm for calculating the Weibull maximum likelihood estimates were the same as described in Engeman and Keefe (1982). The criteria we selected for comparing the performance of the density estimators is the relative root mean squared error (RRMSE) which is calculated by

$$\text{RRMSE} = \left(\frac{\Sigma(\hat{D} - D)^2}{I} \Big/ D^2\right)^{1/2} \tag{19}$$

where $D$ is the true density, $\hat{D}$ is its estimate, and $I$ is the number of iterations in the simulation.

At each iteration of each simulation a new population of objects was defined. Density was estimated without movement using the Fourier series estimator and the Weibull estimator. Movement was then included and the density estimated again by both methods. Each simulation run was comprised of 1000 iterations (generally much larger than most line transect simulation studies), which provided good precision as indicated by RRMSE results remaining almost constant over the movement levels in each density by parameter sighting combination.

Forty-five simulation runs were made. For the sighting function in Eq. (16), all combinations of the three densities by the four movement parameters by the three sighting parameters were simulated (36 runs). Nine additional simulation runs were made for the sighting function in Eq. (17). These runs only considered the moderate movement situation ($p_3 = 1$) for all combinations of the three densities by the three sighting parameters.

Table 1 RRMSE results when using the sighting function in Eq. 16

| Movement parameter | Sighting parameter | Density = 2 | | | | Density = 4 | | | | Density = 8 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Without movement | | With movement | | Without movement | | With movement | | Without movement | | With movement | |
| | | W* | F* | W | F | W | F | W | F | W | F | W | F |
| | 0.5 | 0.49 | 0.36 | 0.91 | 1.04 | 0.48 | 0.28 | 0.91 | 1.05 | 0.41 | 0.21 | 0.92 | 1.03 |
| 0.25 | 1.5 | 0.25 | 0.33 | 0.69 | 1.02 | 0.22 | 0.22 | 0.68 | 1.00 | 0.20 | 0.18 | 0.67 | 1.02 |
| | 3.0 | 0.18 | 0.30 | 0.57 | 1.16 | 0.17 | 0.22 | 0.58 | 1.12 | 0.17 | 0.14 | 0.58 | 1.07 |
| | 0.5 | 0.49 | 0.40 | 0.86 | 1.10 | 0.48 | 0.29 | 0.87 | 1.10 | 0.43 | 0.21 | 0.87 | 1.09 |
| 0.50 | 1.5 | 0.25 | 0.32 | 0.55 | 1.09 | 0.22 | 0.22 | 0.52 | 1.16 | 0.21 | 0.18 | 0.49 | 1.20 |
| | 3.0 | 0.18 | 0.29 | 0.41 | 1.10 | 0.17 | 0.20 | 0.41 | 1.15 | 0.17 | 0.16 | 0.41 | 1.24 |
| | 0.5 | 0.49 | 0.38 | 0.68 | 0.72 | 0.49 | 0.29 | 0.69 | 0.77 | 0.42 | 0.23 | 0.67 | 0.90 |
| 1.00 | 1.5 | 0.25 | 0.32 | 0.36 | 1.15 | 0.22 | 0.21 | 0.33 | 1.16 | 0.20 | 0.20 | 0.29 | 1.16 |
| | 3.0 | 0.17 | 0.31 | 0.26 | 1.15 | 0.17 | 0.22 | 0.26 | 1.17 | 0.17 | 0.16 | 0.26 | 1.17 |
| | 0.5 | 0.49 | 0.37 | 0.57 | 0.48 | 0.49 | 0.29 | 0.57 | 0.58 | 0.42 | 0.22 | 0.51 | 0.96 |
| 2.00 | 1.5 | 0.25 | 0.33 | 0.29 | 0.55 | 0.23 | 0.23 | 0.25 | 0.68 | 0.21 | 0.17 | 0.22 | 0.92 |
| | 3.0 | 0.17 | 0.30 | 0.20 | 0.64 | 0.17 | 0.22 | 0.20 | 0.75 | 0.17 | 0.15 | 0.26 | 1.17 |

*W and F signify results from the Weibull and Fourier series estimators, respectively.

### *4.3 Simulation Results*

The RRMSE results from the 36 simulation runs for the sighting function in Eq. (16) are given in Table 1. As would be expected, the RRMSEs calculated prior to the addition of movement in each simulation remained nearly constant over the four movement levels in each density by sighting parameter combination (e.g., consider the without movement column under Density = 2 and note that each time the sighting parameter is 0.5, the Weibull RRMSE is 0.49). Some of the results from the no movement situations are noteworthy. For the sighting parameter of $p_1 = 0.5$, the RRMSE for the Fourier series estimates ranged from 0.21 to 0.40 whereas the RRMSE from the Weibull estimates were always higher, ranging from 0.41 to 0.49. However as the sighting parameter $p_1$ was increased to a value of 3.0, the RRMSE for the Weibull estimate approached that of the Fourier series (ranging from 0.17 to 0.18 versus 0.14 to 0.31 for the Fourier series) with the Weibull estimator's RRMSE surpassing that of the Fourier series as the density decreased (ranging from 0.17 to 0.18 versus 0.29 to 0.31 for the Fourier series when the density equalled 2). This indicates that for the situation when no movement occurs, the Weibull estimator may perform adequately, at least in low density situations.

When movement (even light movement) was incorporated the Fourier series estimator's RRMSE dramatically increased. The RRMSEs for the Weibull estimator also increased when movement was added; however, the increase was not as dramatic. In many cases the RRMSE for the Fourier series estimator were over 3 or 4 times as great as that for the Weibull estimator. As a gross overview to the simulation results, the mean RRMSEs over all simulation runs when there is no movement was 0.29 for the Weibull estimator versus 0.25 for the Fourier series estimator. With the addition of movement the overall mean RRMSE for the

**Table 2** Analysis of variance of observed RRMSE values in Table 1 where estimation procedure is the fourth factor

| *Source of variation* | *Degrees of freedom* | *RRMSE without movement* | | *RRMSE with movement* | |
|---|---|---|---|---|---|
| | | *Mean square* | *F-ratio* | *Mean square* | *F-ratio* |
| Movement parameter (MP) | 3 | 0.00008 | 0.97 | 0.45501 | 398.54 |
| Sighting parameter (SP) | 2 | 0.22331 | 2600.26 | 0.08569 | 75.06 |
| MP × SP | 6 | 0.00007 | 0.82 | 0.01198 | 10.49 |
| Density (D) | 2 | 0.05438 | 633.27 | 0.02613 | 22.89 |
| MP × D | 6 | 0.00008 | 0.92 | 0.01833 | 16.06 |
| SP × D | 4 | 0.00133 | 15.52 | 0.00158 | 1.38 |
| MP × SP × D | 12 | 0.00004 | 0.41 | 0.00149 | 1.30 |
| Estimator (E) | 1 | 0.02136 | 248.67 | 4.03753 | 3536.53 |
| MP × E | 3 | 0.00006 | 0.73 | 0.07311 | 64.04 |
| SP × E | 2 | 0.08546 | 995.11 | 0.50451 | 441.90 |
| D × E | 2 | 0.01948 | 226.88 | 0.04097 | 35.88 |
| MP × SP × E | 6 | 0.00004 | 0.46 | 0.01875 | 16.42 |
| MP × D × E | 6 | 0.00005 | 0.62 | 0.02026 | 17.75 |
| SP × D × E | 4 | 0.00057 | 6.64 | 0.00054 | 0.47 |
| Error | 12 | 0.00009 | | 0.00114 | |

Weibull is 0.52 which is substantially lower than that for the corresponding Fourier series mean of 1.00. As density and/or sighting parameters ($p_1$) increased, the RRMSE for both estimators improved (decreased). However, that for the Weibull estimator improved more rapidly and to a larger extent than that for the Fourier series estimator.

Although inspection of Table 1 indicates patterns of relative performance for the two estimation methods, some statistical analyses of the RRMSE results seem in order. A four-factor factorial analysis of variance was performed for the movement and no movement situations. The four factors in the analyses were estimation procedure, density, movement parameter and sighting parameter.

Inspection of the first set of mean squares in Table 2 reveals that the sighting parameter contributes by far the most to total variation when there is no movement. Also important, in descending order of mean squares, are the interaction between sighting parameter and estimator, density, estimator and the density-estimator interaction. This suggests, as indicated in the above discussion, that without movement the key to adequate estimation lies in ability to detect the objects more than anything else. The density at which the objects actually occur also affects the ability to estimate density. Estimation method and its interaction with both of these factors affect the density estimates.

In contrast with the first half of Table 2, the second half, which investigates the effects of movement, indicates that the most important contribution to density estimation is the method of estimation. The effect due to the estimation method has a mean square larger than the other 13 effects combined. Next most important (distantly) is the sighting parameter by estimation method interaction followed

**Table 3** RRMSE results when using the sighting function in Eq. 17 and the movement parameter is fixed at 1.00

| | Sighting parameter | Density = 2 | | Density = 4 | | Density = 8 | |
|---|---|---|---|---|---|---|---|
| | | Weibull | Fourier | Weibull | Fourier | Weibull | Fourier |
| Without movement | 0.50 | 0.61 | 0.41 | 0.63 | 0.35 | 0.61 | 0.28 |
| | 0.75 | 0.49 | 0.35 | 0.50 | 0.30 | 0.51 | 0.25 |
| | 1.00 | 0.40 | 0.30 | 0.44 | 0.27 | 0.45 | 0.22 |
| With movement | 0.50 | 0.77 | 0.84 | 0.78 | 0.93 | 0.77 | 1.06 |
| | 0.75 | 0.63 | 0.95 | 0.64 | 1.08 | 0.64 | 1.09 |
| | 1.00 | 0.53 | 1.05 | 0.56 | 1.10 | 0.57 | 1.10 |

**Table 4** Analysis of variance of observed RRMSE values in Table 3 where estimation procedure is the third factor

| Source of variation | Degrees of freedom | RRMSE without movement | | RRMSE with movement | |
|---|---|---|---|---|---|
| | | Mean square | F-ratio | Mean square | F-ratio |
| Sighting parameter (SP) | 2 | 0.02774 | 1997.20 | 0.00240 | 1.42 |
| Density (D) | 2 | 0.00254 | 182.80 | 0.00927 | 5.49 |
| SP × D | 4 | 0.00035 | 25.00 | 0.00077 | 0.45 |
| Estimator (E) | 1 | 0.20267 | 14 592.40 | 0.60867 | 360.40 |
| SP × E | 2 | 0.00417 | 300.40 | 0.05002 | 29.62 |
| D × E | 2 | 0.00604 | 434.80 | 0.00549 | 3.25 |
| Error | 4 | 0.00001 | | 0.00169 | |

closely by the movement parameter. The $F$-ratios are also presented for both halves of Table 2 for the reader's information, although the theoretical justification for their use on RRMSEs is not available nor suggested.

Table 3 contains the RRMSE results from the 9 simulation runs that used the sighting function given in Eq. (17). Only the moderate movement ($p_3 = 1$) situation was considered. With movement, the relative comparisons between the Weibull and Fourier series estimators remain the same as for the simulations using the sighting function from Eq. (16), where the Weibull estimator appears substantially superior to the Fourier series estimator. However, when there is no movement, the Fourier series estimator is consistently superior to the Weibull estimator. With movement the overall mean RRMSE for the Weibull estimator is 0.65 versus 1.02 for the Fourier series estimator. Without movement the overall mean for the Weibull is 0.52 and that for the Fourier series is 0.30.

A three-factor factorial analysis of variance (only one level of movement was considered) was used to study the movement and no movement RRMSE results in Table 3. Inspection of Table 4 indicates that in both the without and with movement cases, estimation method contributes the most to total variation.

Sighting ability also seems to contribute when there is no movement, but not when there is movement.

## CONCLUSIONS

The results indicate that the Weibull method is superior to the Fourier series estimator, even when only light movement is involved. However, the Fourier series estimator is known to have problems when there is movement (Burnham *et al.*, 1980). Additionally, the Weibull estimator had superior performances for some (smaller) densities and sighting function parameterizations where no movement is involved. The Weibull estimator consistently outperforms the Fourier series estimator when movement is involved.

The purpose of this paper was to present a new approach to density estimation from line transect data where the animals move away from the line in response to the observer. We feel that this approach shows promise and we hope that it can provide the groundwork for producing an estimation procedure that works well for the all too frequently encountered situation of animals moving away from the transect line in response to the observer and also when no such movement exists.

### *Acknowledgements*

The writers thank Dave Otis, Ken Burnham and an anonymous referee for their helpful reviews.

### *References*

Burnham, K. P., Anderson, D. R. and Laake, J. L. (1980). Estimation of density from line transect sampling of biological populations. *Wildlife Monographs* **72**, 1–202.

Crain, B. R., Burnham, K. P., Anderson, D. R. and Laake, J. L. (1978). A Fourier series estimator of population-density for line transect sampling. *Utah St. Univ. Press, Logan, Utah.* 25 pp.

Engeman, R. M. and Keefe, T. J. (1982). On generalized least squares estimation of the Weibull distribution. *Communications in Statistics, Theory and Methods* **11**, 2181–2193.

Engeman, R. M. and Keefe, T. J. (1985). Two-step estimators of the scale parameters of the Weibull distribution. *Computers and Biomedical Research* **18**, 391–396.

Gates, C. E., Marshall, W. H. and Olson, D. P. (1968). Line transect method of estimating grouse population densities. *Biometrics* **24**, 135–145.

Laake, J. L., Burnham, K. P. and Anderson, D. R. (1979), User's manual for program TRANSECT. *Utah St. Univ. Press, Logan, Utah.* 26 pp.

Pollock, K. H. (1978). A family of density estimators for line transect sampling. *Biometrics* **34**, 475–478.

Ramsey, F. L. (1979). Parametric models for line transect surveys. *Biometrika* **66**, 505–512.

Seber, G. A. F. (1986). A review of estimating animal abundance. *Biometrics* **42**, 267–292.